基于社交媒体数据的心理指标识别建模：机器学习的方法

苏悦[1,2]　刘明明[1,3]　赵楠[1]　刘晓倩[1]　朱廷劭[1,2]

([1] 中国科学院心理研究所，北京 100101)
([2] 中国科学院大学心理学系，北京 100049)
([3] 联想研究院，北京 100094)

摘　要　心理指标识别建模是基于海量数据结合计算机机器学习算法识别心理特征的一种新兴方式。由于传统纸笔测量方式所存在的诸多限制，本文对基于社交媒体数据的心理建模方法及应用于心理测量的可行性进行综述，介绍了特征及提取方法、常用机器学习算法以及应用场景，并对心理指标识别建模的优势和不足进行了总结与展望。该测量方法基于社交媒体数据，相比自我报告法具有时效性高、可回溯测量、生态效度好等独特优势。然而，基于社交媒体的心理指标识别建模方法也在学习成本、硬件成本等方面存在局限性。未来研究人员需要进一步探索社会媒体信息与用户心理变量间的关联机制，并将心理指标识别模型同传统心理学研究方法结合进行更多的探索和应用。心理指标识别建模结合心理测量基本原理和计算机领域机器学习的技术，将为心理学研究打开一扇新的大门。

关键词　心理建模，心理测量，社交媒体，机器学习，心理预测模型

利用社交媒体数据建立心理指标的识别模型，是通过将被研究者心理测试的自我报告结果与其社交媒体数据相结合，采用机器学习的方法建立两者之间的映射，从而可以实现通过分析用户的社交媒体行为数据直接完成对其心理特征高准确度的自动识别。用户在社交网络上的在线行为数据为心理指标识别建模(下文简称“心理建模”)的开展提供了便于获取的海量行为数据。利用社交媒体数据进行心理建模成为了心理测量的一种新兴方法。

收稿日期：2020-01-14
* 国家社科基金重点项目(17AZD041)；国家自然科学基金项目(31700984)；中国科学院青年创新促进会资助。
苏悦和刘明明对本文的贡献相当，两人为共同第一作者。
通信作者：朱廷劭, email: tszhu@psych.ac.cn

目前心理测量领域应用最为广泛的方法是自我报告法(Robins et al., 2007)。自我报告法能够提供丰富的信息，为研究提供宝贵的自我视角，同时因其操作简便而被广泛使用(Paulhus & Vazire, 2007)。然而自我报告法可能存在以下问题(Dunning et al., 2005)：首先，由于人记忆的固有局限性，当研究人员在用自我报告法进行回溯性研究时，通常很难做到与过去时间点精确匹配的测量；其次，自我报告法受制于测量的人力物力，且涉及到问卷的填写、回收和处理，整个流程耗时长，时效性较差，因而难以大规模地进行频率较高的测量；最后，自我报告法依赖于被试的主动配合，当被试不愿配合，或者不适宜给被试增加额外负担时，自我报告法往往难以顺利开展。

近年来，随着互联网的普及，社交媒体逐渐成为人们生活的重要组成部分。用户的在线行为能够通过电子记录在网络空间中被实时保存下来，形成自然情境下丰富的用户行为数据，为心理测量提供了新的数据平台和研究途径。许多研究结果证明，用户在社交媒体上的行为数据蕴含了大量的心理学含义，是了解人们认知、情感、人格、心理健康等心理过程的另一扇窗。例如，在社交媒体上的浏览时长与用户的社交意愿正相关，社交网站上的好友数量与用户的害羞程度负相关(Orr et al., 2009)。Gosling 等人(2011)认为用户的好友数量、发状态频率等 Facebook 上的行为数据分别与大五人格的五个维度存在显著相关，这表明有可能利用网络数据对用户的人格进行估计。除行为特征外，用户在社交媒体上发布的文本信息(Qiu et al., 2012)、文字表情(Park et al., 2015)等均被发现与心理特征存在显著关联，其效应量在中等以上(Carvalho & Pianowski, 2017)，这表明利用社交媒体数据建立识别心理指标的计算模型具有可行性。

基于以上研究结果，不少学者开展了借助社交媒体数据进行心理建模的研究。本文在梳理心理建模方法的基础上，对心理建模作为心理测量方式的可行性及有效性进行讨论分析，并对其未来应用领域和发展趋势进行展望。

1 心理建模的一般过程

心理建模的一般过程包括社交媒体数据准备(Data Collection)、特征提取(Feature Extraction)、特征选择(Feature Selection)、数据建模(Modeling)、交叉验证(Cross-Validation)以及结果输出(Output)几个主要部分。一般建模过程如图 1 所示。

在心理建模时，首先需要获得用户的社交媒体数据以及对应的心理特征自我报告评分作为心理模型的效标。目前研究中用到的自我报告评分多为自评量表分数，另有些心理模型的效标采用客观指标例如职业阶级(Preoţiuc-Pietro, Lampos, & Aletras, 2015)、收入水平

(Preoţiuc-Pietro, Volkeva, et al., 2015)等。利用社交媒体数据进行心理建模的用户数量一般较大，例如在 Facebook 应用 MyPersonality 上开展的研究用户数量一般在 1000 名以上，最多可达 39 万名用户(He et al., 2014)。

其次，对于社交媒体数据进行量化编码，即特征提取步骤。常用的编码方式包括分类、频数、频率、建立稀疏矩阵等。社交媒体行为信息，例如微博数量，可进行频数的统计和分析(Farnadi et al., 2013; Hao et al., 2014)；文本信息可利用已有的词典进行词频统计(Eichstaedt et al., 2015; Lampos et al., 2014)；社交信息可用于建立反映用户与关注、点赞以及转发用户之间关系的社交矩阵，以进行进一步分析(De Choudhury et al., 2014; Gittelman et al., 2015)。

再次，选用适当的机器学习方法将用户的自我报告得分与相应的社交媒体语言行为特征建立映射关系，并且采用交叉检验，验证模型的计算效果。交叉检验是机器学习建模过程中最为常用的模型性能评估方法，具体操作是将数据集分为训练集和测试集，并用训练集建模、用测试集评估模型性能，数据集将划分多次直至每个数据都做过训练集也做过测试集。交叉验证能够充分利用原始数据并避免了随机划分不均衡对模型性能的影响，同时也可以尽量避免模型的过拟合。

最后，得到基于社交媒体数据的心理特征识别模型。当输入同质用户的社交媒体数据时，心理模型可根据模型特点自动进行特征提取、模型计算并输出用户的心理特征值。

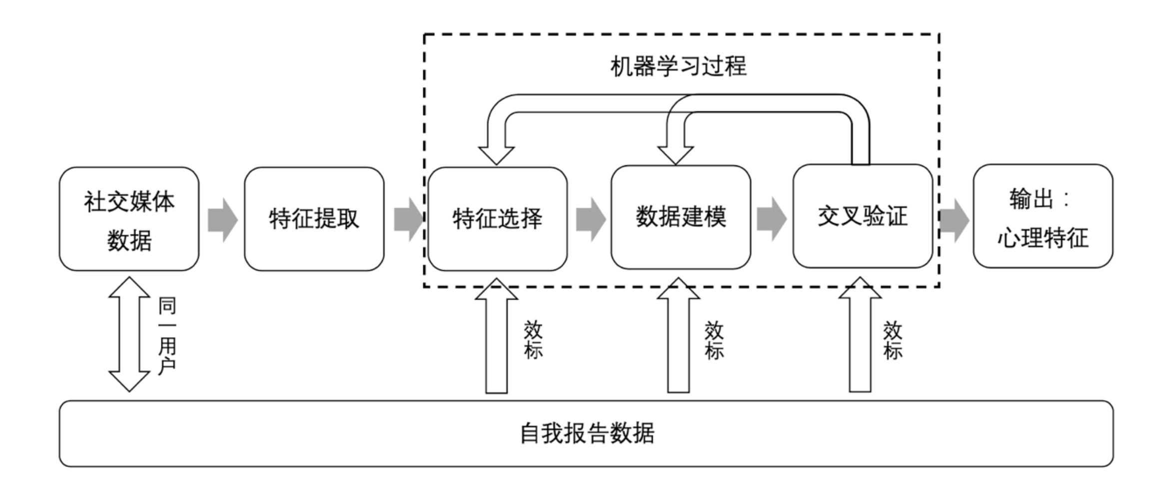


图 1 心理建模的一般过程示意图

2 社交媒体数据类型

由于社交媒体平台本身所包含的多样性信息和丰富的功能，其也给研究人员提供了多种类型的数据。在研究中，人们往往仅选取一种或者几种数据类型进行深入分析和应用。社交媒体数据根据记录形式的不同主要可以分为个人账户信息和使用信息、文本信息、社交网络

信息、图片信息以及其他线索。在心理建模的过程中，不同类别的社交媒体数据可以通过不同的量化编码方式，从丰富的电子记录中提取出若干的数据特征。

2.1 个人账户信息和使用信息

个人账户信息包括用户昵称、性别、生日、居住地、自我简介、隐私设置、头像信息、个性化设置 (郑敬华 等, 2018; Bai et al., 2014; Gao et al., 2013) 等与该社交媒体账户有关的基础信息。区别于自我报告法获得的人口统计学信息，社交媒体上的个人账户信息往往选择性填写，造成缺失数据较多且与真实情况有所出入。因此，在利用个人账户信息进行建模时需要注意个人账户信息与人口学统计信息的区分 (Markovikj et al., 2013)。

社交媒体的使用信息指的是用户在使用社交媒体时留有的大量电子浏览记录，许多研究使用这些“电子足迹”对个体进行心理建模，它们包括发贴的时间(De Choudhury et al., 2013)、首次登录的时间(Nie et al., 2014)、在线时长(Bai et al., 2012)、帖子总数(Celli et al., 2013)、帖子包含 URL 数量(Adalı & Golbeck, 2014)等。

2.2 文本信息

文本作为社交媒体的主要呈现内容之一，也成为了目前研究人员进行心理建模研究使用最为广泛的数据类型，目前已发展出了许多成熟的文本处理技术，包括词频统计、词向量构造、主题模型、自建词典等。

在进行词频统计时，tf-idf 算法可以过滤掉常见的词语，保留重要的词语(Salton & Buckley, 1988)，目前广泛地应用于心理建模的相关研究中(Peng et al., 2015; Seneviratne et al., 2015)。在构造词向量过程中，将具体的词信息转换为词向量的常用转换算法包括 n-gram(王晶晶 等, 2014; Brown et al., 1992; Kern et al., 2014; Mohammad & Kiritchenko, 2015)，Word2Vec (张璞 等, 2019; Rong, 2014; Garten et al., 2016) 和 GloVe (Pennington et al., 2014; Arnoux et al., 2017)。对于主题模型，曹奔等人(2018)针对其在心理学文本分析领域的应用进行了细致的阐述。主题模型也广泛地应用于人格预测(Hu et al., 2017; Liu, Y. Z. et al., 2016)与心理健康分析(Smith et al., 2018; Zhang et al., 2014) 的模型中。在自建词典方面，研究者们编制了许多以情绪、认知、社会关系等为主题的词典，包括 LIWC(Linguistic Inquiry and Word Count)心理语义词典(Pennebaker et al., 2007)、MRC 心理语言数据库(Wilson, 1988)、NRC 情绪词典(Mohammad & Turney, 2013)、ANEW 情绪词典(Nielsen, 2011)、PMI (pointwise mutual information)文字表情词典(Park et al., 2015)、Ekman 情感词典(Volkova & Bachrach, 2015)、道

德基础词典(Moral Foundations Dictionary, MFD) (Haidt et al., 2009)、NLTK 词典 (the Natural Language Toolkit; Loper & Bird, 2002)、Afinn 词典(详见：http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010)、H4Lvd 词典 (详见：http://www.wjh.harvard.edu/~inquirer/inqdict.txt)等。许多基于社交媒体的心理建模研究利用自建词典取得了良好的计算效果 (李昂 等, 2015; 娜迪热, 胡俊, 2018; Golbeck, Robles, Edmondson, & Turner, 2011; Mairesse et al., 2007)。

2.3 社交网络信息

社交网络信息指在社交媒体上用户与其他用户互动的信息。目前的心理建模研究常用的社交网络特征包括关注、点赞、转发、提及、评论等。其中，点赞数据在心理建模中使用得最为广泛。通过提取用户的点赞列表，可以构建用户-主题稀疏矩阵，将每个点赞主题作为一个用户特征列，以此特征列作为自变量进行计算分析(Youyou et al., 2015; Praet et al., 2018)。个体的关注、点赞、转发、提及、评论等社交行为可以构成自我中心的社交网络(ego network)，从而可以计算各个网络的社交网络指标，包括网络大小(白朔天 等, 2014; Bachrach et al., 2012)、网络密度(Celli et al., 2013; Kosinski et al., 2014)、出度和入度(Hao et al., 2014; Li et al., 2014)、中心度(centrality)和集聚系数(transitivity)(Golbeck, Robles, & Turner, 2011; Markovikj et al., 2013)等指标都可以进行建模计算。进一步，可以通过用户的互动情况对其影响力进行分析。常用的用户影响力指标包括“Klout”与“TIME”两项(Sumner et al., 2012; Lima & de Castro, 2014)。

2.4 其他信息

除上述数据类型外，图片信息也在心理建模分析中被广泛使用，包括根据图片的色彩、图像组成、内容特性、图像的亮度等进行心理建模 (Liu, L. et al., 2016; Segalin et al., 2017; Skowron et al., 2016; You et al., 2014)。

此外，近年来有越来越多的社交媒体数据被证明可以用于心理建模，如谷歌应用(Google play app)(Seneviratne et al., 2014)、Facebook 开放的多个应用程序接口(Application Programming Interface，API)(Annalyn et al., 2018; Saha et al., 2017)等。

随着移动互联网的普及，越来越多的移动特征可以被提取与利用，例如大众点评中用户点评的餐厅地点及消费类型(Zhong et al., 2015)。Kalimeri 等人(2019)利用手机浏览网络信息、

手机应用信息与网页浏览信息建立道德基础与人类价值计算模型，计算准确率可达 0.6~0.7，利用手机数据建立的模型相关与网页数据效果相当。

3 心理建模的常用方法

目前，利用社交媒体数据的心理建模方法根据输出类型主要分为分类模型(Classification model)与回归模型(Regression model)。在心理建模时，往往不会只选用一个建模方法，而是尝试在同一数据集上采用多个方法进行训练，通过不断调整和比较，最终确定一个或多个最优模型。

分类模型是指利用社交媒体数据将用户根据一定方法分为两类或多类。分类的结果可以是二分变量，例如性别；也可以人为将结果划分为二分变量，例如将大五人格外向性得分的平均数作为分割点根据分数高低将人群分为内向和外向两类(Farnadi et al., 2018)。常见的分类模型算法包括逻辑斯蒂回归(Logistics Regression, LR)、邻近算法(k-Nearest Neighbor, kNN)、支持向量机(Support Vector Machines, SVM)、朴素贝叶斯(Naive Bayes, NB)、决策树(Decision Tree, DT)、随机森林(Random Forest, RF)等 (杨剑锋 等, 2019; Singh et al., 2016)。

不同的分类模型算法适用于不同的社交媒体数据。逻辑斯蒂回归的结果易于使用与解释(Peng et al., 2002)，在心理建模中的应用十分广泛，包括政治倾向(Praet et al., 2018)、物质滥用(Kosinski et al., 2013)、人格(Celli et al., 2013)、抑郁(De Choudhury et al., 2014)、自杀意念(De Choudhury et al., 2016)等心理特征的分类。邻近算法较为适合高维社交媒体特征条件下的心理特征计算。此外，支持向量机在分类模型的心理建模中也表现良好(刘宝芹, 牛耘, 2016; Ernala et al., 2019; Hao et al., 2013)。朴素贝叶斯和决策树在处理文本数据和社交媒体使用特征时均具有良好的表现(Bai et al., 2012; Farnadi et al., 2016)。随机森林对特征共线性不敏感，结果对缺失数据和偏态数据的计算较为稳健，可以应用于特征量巨大的计算模型中(Breiman, 2001)。

回归模型是一种利用连续特征及离散特征对连续变量进行计算的模型。心理建模中常用的回归算法包括线性回归(Linear Regression)、套索回归(The Least Absolute Shrinkage and Selection Operator Regression, LASSO regression)、岭回归(Ridge Regression, RR)、高斯过程回归(Gaussian Process Regression，GPR)等。其中，线性回归在心理建模中最为常用(Montgomery et al., 2012)。岭回归常在特征具有多重共线性时使用(Hoerl & Kennard, 1970)，套索回归则是常用于高维和稀疏性特征条件下的线性回归(Hans, 2009)。高斯过程回归可作为一种通用的计算方法，也在社交媒体数据特征下取得了良好的预测效果。

4 心理建模的应用场景

现已有许多研究利用心理建模的方法实现了对多种心理特征的识别，研究内容涵盖了诸多心理学研究场景。

4.1 个人信息预测

利用心理建模方法可以对个人信息进行预测，适用于完善网络用户缺失的个人信息，或对用户进行分类。目前，心理建模可对用户性别(Schwartz et al., 2013)、年龄(Zhong et al., 2015)、是否单身(Li et al., 2014)、职业阶级(Preoţiuc-Pietro, Lampos, & Aletras, 2015)、收入水平(Preoţiuc-Pietro, Volkova, et al., 2015)；是否与父母同居、物质滥用、种族、宗教、党派、性取向、母语、国籍、教育水平、子女性别(Kosinski et al., 2013; Li et al., 2014; Volkova & Bachrach, 2015)；星座、血型(Zhong et al., 2015)等变量进行分类或预测。目前对个人信息预测的建模研究已较为全面，在后续研究中研究人员可以将预测结果作为分类依据进行对比研究，尤其是在某些个人信息不便于获取的场景下较为适用。

4.2 人格判断

研究人员尝试了多种类型的特征对大五人格进行建模计算，包括文本信息、行为信息、多特征结合等。Park 等人(2015)采用海量 Facebook 文本信息构建大五人格计算模型，计算精度在 0.34 到 0.46 之间。还有研究人员利用 Facebook 的点赞情况计算用户的大五人格，计算精度最高可达到 0.47 (Youyou et al., 2015)。Liu 和 Zhu(2016)采用深度学习综合微博文本特征、微博行为特征以及表情符号标签等多重社交媒体特征进行建模计算大五人格，最终精度可达 0.3~0.5。

利用心理模型对人格进行判断可避免主观感受或动机对于测量准确性的影响。Kosinski 等人(2016)认为，机器学习计算的人格甚至比伴侣或好友的判断更为准确。Youyou 等人(2017)利用心理模型进行人格的聚类研究，认为使用心理模型进行人格测量可以统一评价标准，克服用户自评时参考群体的影响以及动机的影响。目前对人格建模的研究已经颇为全面，也涉及到了很多社交媒体特征，于建伟(2018)以及张磊等人(2014)均对近年来基于社交网络的人格分析和人格建模研究做出了详细的综述。但是模型的计算精度还有进一步提升的空间，也需要研究人员进行更多的尝试。

4.3 心理健康状态识别

心理建模可以分析用户在社交媒体上发布的感受与想法，为心理健康状态筛查提供了新途径。许多研究证明，可以利用社交媒体数据识别抑郁(Resnik et al., 2013)、自杀倾向(De Choudhury et al., 2016)、精神分裂(Saha et al., 2017)、人格障碍(Carvalho & Pianowski, 2017)、焦虑水平(Settanni & Marengo, 2015)，甚至身体疾病如心脏病(Mathan et al., 2018)、糖尿病、肥胖(Araujo et al., 2017; Mejova et al., 2018)等身心健康问题。Tsugawa 等人(2015)与 Aldarwish 和 Ahmad(2017)两组研究人员分别基于推特与 Facebook 的文本信息建立抑郁识别模型，分类准确率达到 61%和 63%。Nguyen 等人(2017)利用网络文本信息对抑郁、双相情感障碍、自我伤害、悲伤和自杀群体进行区分，分类准确率最高达 88%。

在利用心理建模对用户进行心理健康筛查后，对于存在心理问题的用户可实现主动推送相关资源等干预措施。Liu 等人(2019)提出线上主动自杀预防(Proactive Suicide Prevention Online，PSPO)的方法，主动识别自杀者并为他们提供有效的心理危机干预资源，提高存在自杀意念的用户的求治行为以及公众的健康意识。目前，这类方法在临床中的应用不断增加，在辅助诊断中也存在巨大的潜力，其更广泛的应用仍需针对更多心理变量的更准确的识别模型的开发，以及在心理健康干预领域不断的人力物力投入。

4.4 政治倾向与舆情监测

随着 Web2.0 时代的到来，社交网络已经成为当今舆情表达的主要场所和快速传达民意的渠道(周阳，2018)。利用社交媒体数据可以实时了解民众的舆论态度以及政治倾向。Praet 等人(2018)根据用户的 Facebook 点赞数据构建美国用户的政治倾向与党派偏好的识别模型，计算效果可达到良好水平。Zhou 等人(2017)通过分析社交媒体数据可实现对实时社会态度以及政治议题的舆情趋势监测。

此外，对于关键心理变量如性别、受教育程度等的识别可进一步实现用户政治倾向与舆论的预估。在美国政治倾向的调查中发现，民主党与共和党选民相比，人格方面外向性与尽责性得分更高；价值导向更倾向于传统，整合性与安全性，同时支持普遍主义价值观(Dirilen-Gümüş, Cross, & Dönmez, 2012)；对于自由主义的态度是区分民主党与共和党选民的重要心理指标，对自由主义态度越积极，越有可能倾向于民主党派(Zschirnt, 2011)。用户的社会态度、人格以及情绪状态都会影响政治事件的舆论走向(周阳，2018)。社交媒体数

据为政治倾向及舆论状态的实时监控提供了可能，可作为网络舆情分析的重要辅助工具，可为维护社会稳定、促进民主开放提供保障。

4.5 基于社交媒体的品牌营销

随着社交媒体的用户数量爆发式增长，越来越多的消费品牌开始关注在社交媒体上的营销途径(Social Media Marketing)。品牌的社交网络账号的粉丝数以及评论直接反应了客户的喜爱程度及品牌流行度(De Vries et al., 2012)。社交媒体上的海量数据可以帮助品牌搜集用户需求(Zhu & Chen, 2015)，定位消费群体(Bolotaeva & Cata, 2010)，确立品牌形象(Walsh et al., 2013)，实现个性化营销(Tucker, 2014)。目前，Facebook 的市场应用已经开放相关 API 接口，研究者可以根据关键词或条件选项轻而易举地定位可能感兴趣的用户(Saha et al., 2017)。Matz 等人(2017)尝试针对不同的人格，给同一个网络广告定制了不同风格的首页图，并根据用户的人格特点进行推送，将广告点击率提升了约 40%。未来的研究中可以进一步通过心理建模的方法利用社交媒体数据预测用户感兴趣的品牌，通过品牌账号的文本信息及评论等信息利用主题模型提取品牌形象关键词，联合真实购买数据与社交媒体行为识别用户的消费心理关键变量，从而制定有针对性的营销策略，将营销效果最大化。

4.6 其他

此外，最新研究不断利用社交媒体数据实现更多样的心理变量识别。He 等人(2014)利用 Facebook 上的帖子内容，建模计算用户的自我控制能力。Lewenberg 等人(2015)构建的识别模型则可利用推特数据聚类计算用户的个人兴趣。Kalimeri 等人(2019)可根据手机或网页浏览信息与时长计算道德基础与人类价值。Dufner 等人(2018)构建社交媒体数据与内隐动机倾向的识别模型。利用心理模型对心理变量进行识别可充分利用社交媒体大数据，探究数据中隐含的更多心理信息，实现更加灵活的实验设计。

5 心理建模为心理测量提供新途径

5.1 心理建模作为新途径的优势

相比于纸笔测量方式，基于社交媒体的心理建模有其独特的应用场景和适用范围，可成为现有测量方式的有益补充，为心理测量提供了新途径。

社交媒体上存在大量具有时间标记的数据，为后续更具生态性的心理变量分析提供了数据基础。自我报告的回答可能存在社会赞许效应，被试受社会期望的影响可能会给出误导性的回答(Gerrig et al., 2010)，对结果的准确性存在一定影响。相比而言，基于社交媒体的心理建模通过分析用户的社交媒体数据可实现对被试的无侵扰式测量，无侵扰式的情境下数据均出自用户的自发行为，代表了用户的真实意愿，因而更具生态效度(朱廷劭 等, 2015)。此外，由于个体在使用网络时处于未受测量的自然状态下，也更有利于对配合性低、社会防御强的人物进行准确的心理刻画(Liu, Xue, et al., 2018)。

心理建模能够提供更为统一的行为测量标准。相比于被试各自进行的主观报告，心理建模测量的是特定平台上的行为，并经由计算机进行统一的特征提取和计算，计算过程一致性高，结果更为客观。

心理建模适用于对大规模人群施测，并且被试覆盖范围可以更广。自我报告法往往选取代表样本进行分析。基于社交媒体的心理建模依靠社交媒体规模巨大的用户群体可以将被试范围尽可能地扩大(朱廷劭 等, 2015)，覆盖不同的地区、职业、性别、年龄。特别是当研究本身针对的对象为社交媒体使用者时，基于社交媒体数据进行心理建模的方法几乎可以覆盖研究对象总体，使结果更为全面和客观，也避免了统计误差对结果的影响。

心理建模可追溯的时间跨度大，可在有记录的任一时间点开展心理学研究。自我报告一般专注于临近时间段进行问卷收集，而社交媒体数据具有时间标记，可以不受限制地对用户各个时间点的心理状态进行回溯，由此可开展横断研究或追踪研究(Kosinski et al., 2013)。在追踪研究中，可以跟踪被试特定时间段内的社交媒体活动从而对某项心理特征进行多次计算。这一方法可以实现快捷的数据收集，避免多次填写问卷的练习效应，尽可能减少在多次实验中被试流失的问题，降低实验误差。

基于社交媒体的心理建模可以有效汇总特定条件的研究对象。我们可以根据话题标签、关键词、共同关注等内容对如兴趣爱好、事件经历、话题讨论等条件下的群体进行汇总，利用社交媒体数据进行心理特征的分析。由于有些研究本身的特殊性，相关被试不易招募。利用心理模型进行心理特征的识别可以提供了解该群体的窗口(Liu, Wu, et al., 2018)。此外，研究对于某话题主动关注的人与令被试被动回答对某话题的关注程度，其思路也是截然不同的。

随着互联网社交媒体的普及，社交媒体数据已经成为记录和洞察人的心理特征与行为规律的重要依据。基于社交媒体数据进行心理特征的分析识别已经具备了可行性且可操

作。为了方便对分析过程中的数据与算法进行比较和选择，表 1 汇总了心理建模中常见的数据特征、机器学习方法及要识别的心理变量的组合，以便研究者在建立心理模型时进行参考。

表 1 心理建模常用特征-场景-算法组合汇总

		应用场景			
		个人信息	人格	心理健康	其他
数据类型	个人账户信息	分类：SVM, GP, **LR**	回归：**M5**、**GPR**、**RR**、**线性回归**、PACE 分类：**SVM**、NB、DT	回归：**LASSO**, **SVR**, **stepwise**	
	文本信息	回归：RR 分类：**SVM**、**GP**、**LR**、NB	回归：**GPR**、**线性回归**、**RR**、**M5**、RFR 分类：**NB**、**SVM**、**ZeroR**、**RF**、**DT**、ZeroR、J48、KNN、LR、NN	回归：**线性回归**、**LASSO**、**SVR**、**stepwise**、PACE 分类：**SVM**、**LR**、NN、RF	回归：RR, GPR [用户影响力] 分类：SVM [情感类别] LR [道德判断，自我监控行为]
	社交网络信息	回归：线性回归、RR 分类：**LR**、SVM、GP	回归：**LASSO**、**GPR**、**线性回归**、**RFR**、**M5**、PACE、RR 分类：**SVM**、**NB**、ZeroR、J48、RF、KNN、LR、NB、DT	回归：**线性回归**、**LASSO**、**SVR**、**stepwise**、PACE 分类：**SVM**、 NN	回归：RR, GPR [用户影响力] 分类：LR [政治倾向]
	社交媒体使用信息		回归：**线性回归**、PACE、GPR 分类：**SVM**、**NB**、DT、J48、RF、ZeroR	回归：线性回归, PACE、LASSO, SVR, stepwise 分类：**SVM**、NN	回归：RR, GPR [用户影响力]
	图片信息	分类：LR, NN	回归：**线性回归**、RFR		
	其他信息	回归：PR，线性回归 分类：**SVM**、**LR**、GP、NB、NN	回归：GPR、线性回归、RFR、LASSO 分类：**NB**、**SVM**、KNN、DT、ZeroR	回归：线性回归	分类：RF [人类价值]

注 1：表中统计的文献为来源于谷歌学术搜索关键词组“social media + predict”，“social media + model”，“social media + machine learning”下 2010 年至 2019 年所有的心理建模原创文章(不包括综述和元分析)共 69 篇。

注 2：表中 GP 指高斯过程分类，RFR 指随机森林回归，NN 指神经网络；表格中加粗字体为以往文献中使用 2 次的算法，加粗加下划线字体为以往研究中使用 3 次以上的算法。

注 3：应用场景中其他类别的[]标明心理模型预测的心理变量。

如表 1 所示，在分类模型中，使用频率最高的算法为支持向量机(SVM)和逻辑回归(LR)等；回归模型中，使用频率最高的算法为线性回归和高斯过程回归(GPR)。这些算法在诸多社交媒体特征的建模中均有使用。研究人员在对于新的心理变量进行建模时，可根据具体应用场景和数据类型，借鉴以上列出的常用建模方法，优先考虑该场景下的常用算法。

值得注意的是，任何一种建模算法都有其适用范围，研究人员在选择算法时需要特别关注算法的前置假设。一味追求暂时的模型效果而忽视算法本身的限制条件反而会阻碍模型的适用范围，需要严格检验数据并对建立的模型不断优化。

为了对心理建模的性能进行评估，依据目标心理特征的属性不同，目前评价标准可划分为两大类。离散型心理指标的主要评价标准有准确率(Accuracy)、精确率(Precision)、接收者操作特征曲线(Receiver Operating Characteristic Curve, 也称 ROC 曲线)、曲线下面积(Area Under the Curve, AUC)等。连续型心理指标的主要评价标准包括预测值与真实值的相关系数(r)、平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)、决定系数(R^2)等。现已有相当多的分类模型的分类准确率可达 80%以上(Celli et al., 2013; Iacobelli et al., 2011; Seneviratne et al., 2014)，相当于心理测量的效标效度计算中总命中率达到了 80%以上；在回归模型中，模型计算结果同真实值之间的相关系数也可以达到 0.3~0.5 之间，相当于通过回归分析对效标效度进行计算并达到了中等相关的水平。综上，不论是哪类心理特征，心理指标识别模型均具有良好的效标效度。

此外也有研究表明心理模型的计算结果具有较高的稳定性和一致性。刘明明(2019)对其所建立的心理模型的重测信度进行了检验，其中，大五人格识别模型的测量间隔为 6 个月，抑郁、自杀可能性、生活满意度识别模型的重测间隔为 1 个月，将重测前后心理特征进行分析，其中大五人格、抑郁、自杀可能性及生活满意度模型的重测信度均在 0.75 以上，达到了较高的可信度，结果如表 2 所示。

表 2 心理模型识别信度检验(刘明明，2019)

计算模型	**重测信度**
大五人格	0.77~0.79
抑郁	0.83
自杀可能性	0.80~0.91
生活满意度	0.84

综上所述，我们发现虽然基于社交媒体的心理建模没有遵循严格的心理量表开发过程，但其模型的计算结果经过信效度检验，模型对心理指标的识别是稳定可靠的。

5.2 心理建模的不足之处

尽管基于社交媒体的心理建模是可行的，并且其计算结果经过一定的信效度检验，但是作为一种新的方法，仍然存在不足之处。

首先，基于社交媒体的心理建模这一新方法具有一定的学习成本。相比于心理学研究人员已经熟练掌握的纸笔测量方式，心理建模过程中涉及到的大量计算机专业知识，相对复杂的心理变量计算过程，都会给心理学领域研究人员对结果的计算和解释带来一定的挑战。

此外，这一新方法还有额外的设备成本。社交网络中的海量用户数据可达 TB 级别，进行处理和分析时，计算机的计算性能和存储性能均将面临更高的要求。

其次，新方法的场景以社交媒体为主，也面临着被试范围受限的问题。虽然社交媒体用户的组成多样，覆盖人群较广，但仍无法覆盖真实世界中的全部用户。社交媒体用户这一特定群体可能会对实验带来群体偏差；不同社交平台的用户也会存在不同的群体特征。这些基于场景的群体偏差对于心理建模研究的影响仍有待进一步探讨。

再次，基于社交媒体的心理建模目前也存在准确度受限的问题。目前，大部分心理变量的识别模型以自评量表得分作为效标进行建模(Kosinski et al., 2015)。模型本身计算的准确程度无法超越纸笔测量，且自评阶段结果的准确性也影响着模型的好坏。目前，已经有许多传统心理测量方法在主观自我报告的基础之上加入客观测量指标，作为主观报告的补充进行聚合指标的研究。例如，睡眠质量可采用睡眠时间，睡眠潜伏期等客观指标进行表示(Devnani & Hegde, 2015)；注意水平可采取主观测量与客观测量结合进行研究，如脑电，眼动水平等(Hopstaken et al., 2016)。未来的心理建模可借鉴客观测量指标，逐步从仅将自我报告得分作为建模效标的计算模式，转变为以结合主观报告和客观测量的综合指标作为建模的目标变量的方式，进一步提升心理模型的内部效度。

最后，目前心理模型的识别精度仍需进一步提高。尽管目前心理分类模型能够达到 0.8~0.9 以上的准确率；回归模型能达到 0.3 以上的中等相关，但仍存在以当前的心理建模精度是否能够达到心理学研究所需良好的信效度等疑问。心理建模的精度提升是基于计算

机与心理测量领域技术发展的缓慢累积过程。目前的计算结果往往是根据概率分布预测的样本分布状态，因此在做临床使用或者针对个体差异化评估时需要谨慎使用(Liu, Xue, et al., 2018)。

6 心理建模的未来发展趋势

利用社交媒体数据进行建模，从而识别心理特征是一种新兴的心理测量方法，在用户个人信息预测、人格判断、心理健康筛查、政治倾向判断以及消费行为预测等方向上具有巨大的潜力。在未来的研究中，有如下发展趋势值得关注：

6.1 社会媒体信息与用户心理变量间的关联机制

利用机器学习算法进行心理建模，可以通过数据特征直接计算出相应的心理指标，然而数据与心理变量间的关联机制却相对难以解释和理解。一些复杂的机器学习算法，例如神经网络或高斯过程，并不直接将某一社交媒体特征与目标心理变量进行关联，而是经过层层变换最终计算出心理变量的离散类别或连续值(Arnoux et al., 2017; Wang & Kosinski, 2018)。另一些建模过程通过降维运算或傅里叶变换转换数据坐标系，从而使得数据特征丧失原有的心理学含义(Praet et al. 2018)。

可解释性强的心理建模方式在心理学研究中往往能提供更多的过程信息，同时也能倒推出更好的心理干预方式。如何将行为数据的处理赋予心理学含义是一项值得关注的话题。一方面研究人员开始尝试解释性强的算法进行建模，如进化模糊系统(Evolutionary Fuzzy Systems)(Fernandez et al., 2019)，新闻处理领域的 I-T-O 算法透明度模型(仇筠茜，陈昌凤, 2018)等。另一方面，越来越丰富的社交媒体数据可以使得应用简便易懂的方式建模便能够取得良好的模型效果(Nave et al., 2018; Kalimeri et al., 2019)。此外，也可以尝试结合心理学理论，如通过心理学现有研究中某些行为特点和心理特质之间的关系对特征进行筛选，从而保证所确定的建模特征在心理学层面的可解释性。

6.2 多数据源特征与多识别模型融合优化模型精度

随着互联网技术与通讯技术的不断发展，越来越多的信息以更加密集的形式呈现在社交媒体上，小视频、3D/4D 影像、虚拟现实等包含大量信息的数据形式高速发展(Roberts &

Foehr, 2008)。心理建模所采用的数据特征从最初单一的文本特征逐渐过渡到图片、手机定位等多元化的综合特征。Azucar 等人(2018)通过元分析证明在进行心理建模时多种特征结合使用的建模效果优于单种特征，多样化特征种类可以更加全面地识别个体心理特征。

新出现的社交媒体数据是否与心理特征存在关联，怎样整合不同类型的特征以取得更优的识别效果，都是亟待研究的新问题。此外，研究表明，融合(fusion)多个机器学习模型往往可以提高整体的识别能力(Yu et al., 2011)。因此，在未来的研究中研究者们可以将这一思路运用到心理特征识别的建模实践中，深入挖掘使用多数据源特征与多识别模型集成的方法，进一步提高模型精度。

6.3 心理模型与传统心理学研究方法的有效结合

基于社交媒体的心理建模方法作为一种补充的测量方式，在自我报告法难以实现的场景下具有一定的优势，因此，将其与传统心理学研究方法结合能够进一步拓展研究范畴。研究人员可以借助这一新方法进行更多的对比实验，如国籍、文化、地域等方面的差异研究。另外，利用心理模型进行测量配合灵活的实验设计可以获取以往传统方式难以获取的样本量和特殊样本群体，更可以突破时间上和被试招募上的限制从而将某些相关性问题推进到因果性角度的探究。

现已有学者利用基于社交媒体的心理建模方法进行实验设计并开展研究。Matz 等人(2017)基于 Facebook 点赞行为对用户的大五人格特点进行识别，并在此基础上进行了定向广告投放，结果表明当广告内容同受众性格相匹配时更易影响他们的行为。该研究涉及了 Facebook 平台上的数百万名用户，传统测量方式无法实现，而基于社交媒体的心理建模这一新方法提供的用户人格特征为后续研究奠定了基础。有些研究针对某些重大生活事件，而这类事件发生后短期内往往难以获得足够的样本进行心理特征的测量，且由于该类事件无法预知从而无法获得前后测的对照。Liu, Xue 等人(2018)基于新浪微博的用户活动利用心理健康模型对于家庭暴力受害者短期内受到的影响进行分析，克服了传统方法无法即时测量的缺陷，研究了受害者在家暴前后短期时间内抑郁、生活满意度等的变化。

6.4 心理建模和脑科学领域的深度融合

心理建模的研究和脑科学的发展相互结合相互促进。一方面，目前的心理建模研究主要集中于对社交媒体用户的网络行为进行分析，通过提取用户的社交媒体行为特征建立心理特征的预测模型。然而在社交媒体行为预测心理特征这一过程中用户行为的心理学机制尤其是脑科学机制尚不明确。从脑科学角度深入挖掘用户的社交媒体行为有助于进一步揭示用户行为背后的神经科学机制，从而增强依据行为特征所建立的心理模型的可解释性。同时在心理建模过程中依据神经科学基础进行特征筛选和提取也有望进一步提高模型的计算性能。另一方面，基于社交媒体的心理建模作为一种交叉领域的方法，可以应用于与认知相关的心理特征的分析研究中。通过这样的方式，我们可以把用户认知相关心理特征、社交媒体互动环境以及大脑的认知活动有效结合在一起，为深入研究个体与个体自然状态下社交活动的心理机制提供可能。社交媒体平台为脑科学的研究提供了用户在线进行社交互动的生态环境，而基于社交媒体的心理建模则能够对用户相应的心理特征进行计算和分析，在此基础上，研究人员可以对人类在网络互动环境下的认知活动等心理过程进行进一步探究。

参考文献

白朔天，袁莎，程立，朱廷劭. (2014). 多任务回归在社交媒体挖掘中的应用. *哈尔滨工业大学学报*, *46*(09), 100–104+110.

曹奔，夏勉，任志洪，林秀彬，徐升，赖丽足，王琪，江光荣. (2018). 大数据时代心理学文本分析技术——“主题模型”的应用. *心理科学进展*, *26*(05), 770–780.

仇筠茜，陈昌凤. (2018). 基于人工智能与算法新闻透明度的“黑箱”打开方式选择. *郑州大学学报（哲学社会科学版)*, *51*(5), 84–88.

李昂，郝碧波，白朔天，朱廷劭. (2015). 基于网络数据分析的心理计算：针对心理健康状态与主观幸福感. *科学通报*, *60*(11), 994–1001.

刘宝芹，牛耘. (2016). 基于情绪特征的中文微博用户性别识别. *计算机工程与科学*, *38*(09), 1917–1923.

刘明明.(2019).*基于社交媒体数据的心理特征自动识别新方法研究*(硕士学位论文).中国科学院大学,北京.

娜迪热，胡俊. (2018). 基于用户社交网络数据的人格倾向性分析及预测模型的建立. *电脑知识与技术*, *14*(07), 6–11.

王晶晶，李寿山，黄磊. (2014). 中文微博用户性别分类方法研究. *中文信息学报*, *28*(6), 150–155.

杨剑锋，乔佩蕊，李永梅，王宁. (2019). 机器学习分类问题及算法研究综述. *统计与决策*, *35*(06), 36–40.

于建伟. (2018). 基于社交网络的人格分析与预测. *现代计算机(专业版)*, (04), 29–34.

张磊，陈贞翔，杨波. (2014). 社交网络用户的人格分析与预测. *计算机学报*, *37*(08), 1877–1894.

张璞，陈超，陈韬，王永. (2019). 两分类器融合的中文微博用户性别分类方法. *计算机工程与设计*, *40*(01), 276–280.

郑敬华，郭世泽，高梁，赵楠. (2018). 基于多任务学习的大五人格预测. *中国科学院大学学报*, *35*(04), 550–560.

周阳. (2018). *基于社会媒体大数据的舆情预警*(博士学位论文).中国科学院大学,北京.

朱廷劭，汪静莹，赵楠，刘晓倩. (2015). 论大数据时代的心理学研究变革. *新疆师范大学学报(哲学社会科学版)*, (04), 2+108–115.

Adalı, S., & Golbeck, J. (2014). Predicting personality with social behavior: a comparative study. *Social Network Analysis and Mining*, *4*(1), 159.

Aldarwish, M. M., & Ahmad, H. F. (2017, March). Predicting depression levels using social media posts. *In 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)* (pp. 277–280). IEEE.

Annalyn, N., Bos, M. W., Sigal, L., & Li, B. (2018). Predicting personality from book preferences with user-generated content labels. *IEEE Transactions on Affective Computing*.

Araujo, M., Mejova, Y., Weber, I., & Benevenuto, F. (2017, June). Using Facebook ads audiences for global lifestyle disease surveillance: Promises and limitations. In *Proceedings of the 2017 ACM on Web Science Conference* (pp. 253–257). ACM.

Arnoux, P. H., Xu, A., Boyette, N., Mahmud, J., Akkiraju, R., & Sinha, V. (2017, May). 25 Tweets to Know You: A New Model to Predict Personality with Social Media. In *Eleventh International AAAI Conference on Web and Social Media*.

Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, *124*, 150–159.

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012, June). Personality and patterns of Facebook usage. *In Proceedings of the 4th Annual ACM Web Science Conference* (pp. 24–32). ACM.

Bai, S., Yuan, S., Hao, B., & Zhu, T. (2014). Predicting personality traits of microblog users. *Web Intelligence and Agent Systems: An International Journal*, *12*(3), 249–265.

Bai, S., Zhu, T., & Cheng, L. (2012). Big-five personality prediction based on user behaviors at social network sites. *arXiv preprint arXiv:1204.4809*.

Bolotaeva, V., & Cata, T. (2010). Marketing opportunities with social networks. *Journal of Internet Social Networking and Virtual Communities*, *2010*, 1–8.

Breiman, L. (2001). Random forests. *Machine Learning, 45(1)*, 5–32.

Brown, P. F., Desouza, P. V., Mercer, R. L., Pietra, V. J. D., & Lai, J. C. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, *18*(4), 467–479.

Carvalho, L. de F., & Pianowski, G. (2017). Pathological personality traits assessment using Facebook: Systematic review and meta-analyses. *Computers in Human Behavior*, *71*, 307–317.

Celli, F., Pianesi, F., Stillwell, D., & Kosinski, M. (2013). Workshop on Computational Personality Recognition: Shared Task. *Proceedings of the Workshop on Personality Recognition*, *2006*, 2–5.

De Choudhury, M., Counts, S., & Horvitz, E. (2013, May). Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 47–56). ACM.

De Choudhury, M., Counts, S., Horvitz, E. J., & Hoff, A. (2014, February). Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 626–638). ACM.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016, May). Discovering shifts to suicidal ideation from mental health content in social media. *In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2098–2110). ACM.

De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, *26*(2), 83–91.

Devnani, P. A., & Hegde, A. U. (2015). Autism and sleep disorders. *Journal of Pediatric Neurosciences*, *10*(4), 304–307.

Dirilen‐Gümüş, Ö., Cross, S. E., & Dönmez, A. (2012). Who Voted for Whom? Comparing Supporters of Obama and McCain on Value Types and Personality Traits. *Journal of Applied Social Psychology*, *42*(12), 2879–2900.

Dufner, M., Arslan, R. C., & Denissen, J. J. A. (2018). The unconscious side of Facebook: Do online social network profiles leak cues to users' implicit motive dispositions? *Motivation and Emotion*, *42*(1), 79–89.

Dunning, D., Heath, C., & Suls, J. M. (2005). Picture imperfect. *Scientific American Mind*, *16*(4), 20–27.

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., ... & Weeg, C. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science, 26*(2), 159–169.

Ernala, S. K., Birnbaum, M. L., Candan, K. A., Rizvi, A. F., Sterling, W. A., Kane, J. M., & De Choudhury, M. (2019, April). Methodological gaps in predicting mental health states from social media: Triangulating diagnostic signals. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 134). ACM.

Farnadi, G., Sitaraman, G., Sushmita, S., Celli, F., Kosinski, M., Stillwell, D., ... & De Cock, M. (2016). Computational personality recognition in social media. *User Modeling and User-adapted Interaction*, *26*(2–3), 109–142.

Farnadi, G., Tang, J., De Cock, M., & Moens, M. F. (2018, February). User profiling through deep multimodal fusion. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 171–179). ACM.

Farnadi, G., Zoghbi, S., Moens, M., & Cock, M. De. (2013, June). Recognising Personality Traits Using Facebook Status Updates. *Workshop on Computational Personality Recognition (WCPR13) in International AAAI Conference on Weblogs and Social Media (ICWSM13),* 14–18.

Fernandez, A., Herrera, F., Cordon, O., del Jesus, M. J., & Marcelloni, F. (2019). Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to?. *IEEE Computational Intelligence Magazine*, *14*(1), 69–81.

Gao, R., Hao, B., Bai, S., Li, L., Li, A., & Zhu, T. (2013, October). Improving user profile with personality traits predicted from social media content. In *Proceedings of the 7th ACM Conference on Recommender Systems* (pp. 355–358). ACM.

Garten, J., Boghrati, R., Hoover, J., Johnson, K. M., & Dehghani, M. (2016, July). Morality between the lines: Detecting moral sentiment in text. In *Proceedings of IJCAI 2016 Workshop on Computational Modeling of Attitudes*.

Gerrig, R. J., Zimbardo, P. G., Zimbardo, P. G., Psychologue, E. U., & Zimbardo, P. G. (2012). *Psychology and Life* (20th ed.). Boston: Pearson.

Gittelman, S., Lange, V., Crawford, C. A. G., Okoro, C. A., Lieb, E., Dhingra, S. S., & Trimarchi, E. (2015). A new source of data for public health surveillance: Facebook likes. *Journal of Medical Internet Research, 17*(4), e98.

Golbeck, J., Robles, C., & Turner, K. (2011). Predicting personality with social media. Paper presented at *the CHI '11 Extended Abstracts on Human Factors in Computing Systems*(pp.253–262), ACM. https://doi.org/10.1145/1979742.1979614.

Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011, October). Predicting personality from twitter. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* (pp. 149–156). IEEE.

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking, 14*(9), 483–488.

Haidt, J., Graham, J., & Nosek, B.A. (2009). Liberals and Conservatives Rely on Different Sets of Moral Foundations. *Journal of Personality and Social Inquiry*, *20*(2–3), 110–119.

Hans, C. (2009). Bayesian lasso regression. *Biometrika, 96*(4), 835–845.

Hao, B., Li, L., Gao, R., Li, A., & Zhu, T. (2014, August). Sensing subjective well-being from social media. In *International Conference on Active Media Technology* (pp. 324–335). Springer, Cham.

Hao, B., Li, L., Li, A., & Zhu, T. (2013, July). Predicting Mental Health Status on Social Media A Preliminary Study on Microblog. *15th International Conference on Human-Computer Interaction* 8024, 101–110.

He, Q., Glas, C. A., Kosinski, M., Stillwell, D. J., & Veldkamp, B. P. (2014). Predicting self-monitoring skills using textual posts on Facebook. *Computers in Human Behavior, 33*, 69–78.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. *Technometrics*, *12*(1), 69–82.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A., & Leung, Y. K. (2016). Shifts in attention

during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 878–889.

Hu, Z., Liu, Y. S., Zhang, C. H., & Xu, Y. N. (2017, June). The analysis of topic's personality traits using a new topic model. *In 2017 2nd International Conference on Image, Vision and Computing (ICIVC)* (pp. 1079–1083). IEEE.

Iacobelli, F., Gill, A. J., Nowson, S., & Oberlander, J. (2011). Large Scale Personality Classification of Bloggers. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6975*(PART 2), 487–496. https://doi.org/10.1007/978–3–642–24571–8

Kalimeri, K., Beiró, M. G., Delfino, M., Raleigh, R., & Cattuto, C. (2019). Predicting demographics, moral foundations, and human values from digital behaviours. *Computers in Human Behavior*, *92*, 428–445.

Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., ... & Seligman, M. E. (2014). The online social self: An open vocabulary approach to personality. *Assessment*, *21*(2), 158–169.

Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, *95*(3), 357–380.

Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, *70*(6), 543–548.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–5805.

Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes. *Psychological Methods, 21*(4), 493.

Lampos, V., Aletras, N., Preoţiuc-Pietro, D., & Cohn, T. (2014, January). Predicting and characterising user impact on Twitter. In *14th Conference of the European Chapter of the Association for Computational Linguistics 2014, EACL 2014*(pp. 405–413).

Lewenberg, Y., Bachrach, Y., & Volkova, S. (2015, October). Using emotions to predict user interest areas in online social networks. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1–10). IEEE.

Li, L., Li, A., Hao, B., Guan, Z., & Zhu, T. (2014). Predicting active users' personality based on micro-blogging behaviors. *PloS One*, *9*(1), e84997.

Lima, A. C. E. S., & de Castro, L. N. (2014). A multi-label, semi-supervised classification approach applied to personality prediction in social media. *Neural Networks*, *58*, 122–130.

Liu, L., Preot, D., & Ungar, L. (2016). Analyzing Personality through Social Media Profile Picture Choice. In *Proceedings of the Tenth International AAAI Conference on Web and Social Media(ICWSM)* (pp.211–220). AAAI.

Liu, M., Wu, Y., Jiao, D., Wu, M. S., & Zhu, T. (2018). Literary intelligence analysis of novel protagonists' personality traits and development. *Digital Scholarship in the Humanities*, *34*(1), 221–229.

Liu, M., Xue, J., Zhao, N., Wang, X., Jiao, D., & Zhu, T. (2018). Using social media to explore the consequences of domestic violence on mental health. *Journal of Interpersonal Violence*, 0886260518757756.

Liu, X., & Zhu, T. (2016). Deep learning for constructing microblog behavior representation to identify social media user's personality. *PeerJ Computer Science*, *2*, e81.

Liu, X., Liu, X., Sun, J., Yu, N. X., Sun, B., Li, Q., & Zhu, T. (2019). Proactive Suicide Prevention Online (PSPO): Machine Identification and Crisis Management for Chinese Social Media Users With Suicidal Thoughts and Behaviors. *Journal of Medical Internet Research*, *21*(5), e11705.

Liu, Y. Z., Wang, J. J., & Jiang, Y. C. (2016). PT-LDA: A latent variable model to predict personality traits of social network users. *Neurocomputing*, *210*, 155–163.

Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics* (pp. 63–70). Association for Computational Linguistics.

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research, 30*, 457–500.

Markovikj, D., Gievska, S., Kosinski, M., & Stillwell, D. (2013, June). Mining Facebook data for Predictive Personality Modeling. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 23–26.

Mathan, K., Kumar, P. M., Panchatcharam, P., Manogaran, G., & Varadharajan, R. (2018). A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. *Design Automation for Embedded Systems*, *22*(3), 225–242.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, *114*(48), 12714–12719.

Mejova, Y., Weber, I., & Fernandez-Luque, L. (2018). Online health monitoring using facebook advertisement audience estimates in the United States: Evaluation study. *Journal of Medical Internet Research*, *20*(3), 1–16. https://doi.org/10.2196/publichealth.7217

Mohammad, S. M., & Kiritchenko, S. (2015). Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*, *31*(2), 301–326.

Mohammad, S. M., & Turney, P. D. (2013). Nrc emotion lexicon. National Research Council. http://saifmohammad.com/WebDocs/NRCemotionlexicon.pdf.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). Hoboken, NJ: John Wiley & Sons.

Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., & Rentfrow, J. (2018). Musical preferences predict personality: evidence from active listening and facebook likes. *Psychological Science*, *29*(7), 1145–1158.

Nguyen, T., O'Dea, B., Larsen, M., Phung, D., Venkatesh, S., & Christensen, H. (2017). Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimedia Tools and Applications*, *76*(8), 10653–10676.

Nie, D., Guan, Z., Hao, B., Bai, S., & Zhu, T. (2014, August). Predicting personality on social media with semi-supervised learning. In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)* (Vol. 2, pp. 158–165). IEEE.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In *Proceedings of the ESWC2011 Workshop on Making Sense of Microposts* (pp. 93–98). CEUR Workshop Proceedings.

Orr, E. S., Sisic, M., Ross, C., Simmering, M. G., Arseneault, J. M., & Orr, R. R. (2009). The influence of shyness on the use of Facebook in an undergraduate sample. *CyberPsychology & Behavior*, *12*(3), 337–340.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., ... & Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*(6), 934.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology, 1*, 224–239.

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1), 3–14.

Peng, K. H., Liou, L. H., Chang, C. S., & Lee, D. S. (2015, October). Predicting personality traits of Chinese users based on Facebook wall posts. In *2015 24th Wireless and Optical Communication Conference (WOCC)* (pp. 9–14). IEEE.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. *Austin, TX: liwc. net, 135*.

Pennington, J., Socher, R., & Manning, C. (2014, October). Glove: Global vectors for word representation. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics.

Praet, S., Van Aelst, P., & Martens, D. (2018). I like, therefore I am. Predictive modeling to gain insights in political preference in a multi-party system. University of Antwerp, Faculty of Business and Economics,1–34.

Preoţiuc-Pietro, D., Lampos, V., & Aletras, N. (2015, July). An analysis of the user occupational class through Twitter content. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (Vol. 1, pp. 1754–1764). Association for Computational Linguistics.

Preoţiuc-Pietro, D., Volkova, S., Lampos, V., Bachrach, Y., & Aletras, N. (2015). Studying user income through language, behaviour and affect in social media. *PloS One, 10*(9), e0138717.

Qiu, L., Lin, H., Ramsay, J., & Yang, F. (2012). You are what you tweet: Personality expression and perception on Twitter. *Journal of Research in Personality*, *46*(6), 710–718.

Resnik, P., Garron, A., & Resnik, R. (2013, October). Using topic modeling to improve prediction of neuroticism and depression in college students. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1348–1353). Association for Computational Linguistics.

Roberts, D. F., & Foehr, U. G. (2008). Trends in media use. *The Future of Children*, *18*(1), 11–37.

Robins, R. W., Tracy, J. L., & Sherman, J. W. (2007). What kinds of methods do personality psychologists use?. In R. Robins, C. Fraley, & R. Krueger(Eds.), *Handbook of research methods in personality psychology*(pp.673–678). New York, NY: Guilford Press.

Rong, X. (2014). word2vec parameter learning explained. *arXiv preprint* arXiv:1411.2738.

Saha, K., Weber, I., Birnbaum, M. L., & De Choudhury, M. (2017). Characterizing awareness of schizophrenia among Facebook users by leveraging Facebook advertisement estimates. *Journal of Medical Internet Research, 19*(5), e156.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513–523.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS One, 8*(9), e73791.

Segalin, C., Perina, A., Cristani, M., & Vinciarelli, A. (2017). The Pictures We Like Are Our Image: Continuous Mapping of Favorite Pictures into Self-Assessed and Attributed Personality Traits. *IEEE Transactions on Affective Computing*, *8*(2), 268–285. https://doi.org/10.1109/TAFFC.2016.2516994.

Seneviratne, S., Seneviratne, A., Mohapatra, P., & Mahanti, A. (2014). Predicting user traits from a snapshot of apps installed on a smartphone. *ACM SIGMOBILE Mobile Computing and Communications Review*, *18*(2), 1–8. https://doi.org/10.1145/2636242.2636244

Seneviratne, S., Seneviratne, A., Mohapatra, P., & Mahanti, A. (2015). Your installed apps reveal your gender and more!. *ACM SIGMOBILE Mobile Computing and Communications Review*, *18*(3), 55–61.

Settanni, M., & Marengo, D. (2015). Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. *Frontiers in Psychology*, *6*, 1045.

Shanableh, T., & Assaleh, K. (2010). Feature modeling using polynomial classifiers and stepwise regression. *Neurocomputing*, *73*(10–12), 1752–1759.

Singh, A., Thakur, N., & Sharma, A. (2016, March). A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1310–1315). IEEE.

Skowron, M., Tkalčič, M., Ferwerda, B., & Schedl, M. (2016, April). Fusing social media cues: personality prediction from twitter and instagram. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 107–108). International World Wide Web Conferences Steering Committee.

Smith, R. J., Schwartz, H. A., Preoţiuc-Pietro, D., Eichstaedt, J. C., Asch, D. A., Ungar, L. H., … Crutchley, P. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, *115*(44), 11203–11208. https://doi.org/10.1073/pnas.1802331115

Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012, December). Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. In *2012 11th International Conference on Machine Learning and Applications* (Vol. 2, pp. 386–393). IEEE.

Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., & Ohsaki, H. (2015, April). Recognizing depression from twitter activity. *In Proceedings of the 33rd annual ACM Conference on Human Factors in Computing Systems* (pp. 3187–3196). ACM.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, *51*(5), 546–562.

Volkova, S., & Bachrach, Y. (2015). On Predicting Sociodemographic Traits and Emotions from Communications in Social Networks and Their Implications to Online Self-Disclosure. *Cyberpsychology, Behavior, and Social Networking*, *18*(12), 726–736. https://doi.org/10.1089/cyber.2014.0609

Walsh, P., Clavio, G., Lovell, M. D., & Blaszka, M. (2013). Differences in Event Brand Personality Between Social Media Users and Non-Users. *Sport Marketing Quarterly*, *22*(4), 214–223.

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, *114*(2), 246–257.

Wang, Y., & Witten, I. H. (1999). Pace regression. Department of Computer Science, The University of Waikato, Hamilton, New Zealand.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–10.

You, Q., Bhatia, S., Sun, T., & Luo, J. (2014, December). The eyes of the beholder: Gender prediction using images posted in online social networks. *In 2014 IEEE International Conference on Data Mining Workshop* (pp. 1026–1030). IEEE.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences, 112*(4), 1036–1040.

Youyou, W., Stillwell, D., Schwartz, H. A., & Kosinski, M. (2017). Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends. *Psychological Science*, *28*(3), 276–284.

Yu, S., Tranchevent, L. C., De Moor, B., & Moreau, Y. (2011). *Kernel-based data fusion for machine learning*. Berlin: Springer.

Zhang, L., Huang, X., Liu, T., Li, A., Chen, Z., & Zhu, T. (2014, November). Using linguistic features to estimate suicide probability of Chinese microblog users. In *International Conference on Human Centered Computing* (pp. 549–559). Springer, Cham.

Zhong, Y., Yuan, N. J., Zhong, W., Zhang, F., & Xie, X. (2015). You Are Where You Go: Inferring Demographic Attributes from Location Check-ins. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)* (pp.295–304). ACM.

Zhou, Y., Zhang, L., Liu, X., Zhang, Z., Bai, S., & Zhu, T. (2017). Predicting the trends of social events on Chinese social media. *Cyberpsychology, Behavior, and Social Networking*, *20*(9), 533–539.

Zhu, Y. Q., & Chen, H. G. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business Horizons*, *58*(3), 335–345.

Zschirnt, S. (2011). The origins & meaning of liberal/conservative self-identifications revisited. *Political Behavior, 33*(4), 685–701.

Identifying psychological indexes based on social media data: A machine learning method

SU Yue[1,2], LIU Mingming[1,3], ZHAO Nan[1], LIU Xiaoqian[1], ZHU Tingshao[1,2]

([1] *Institute of Psychology, Chinese Academy of Sciences, Beijing* 100101, *China*)

([2] *Department of Psychology, University of Chinese Academy of Sciences, Beijing* 100049, *China*)

([3] *Lenovo research, Beijing* 100094, *China*)

Abstract: Modeling psychological indexes (i.e., psych-modeling) is an emerging method that uses machine learning algorithms to identify psychological indexes based on big data. This paper reviews the feasibility of psych-modeling methods based on social media data in the field of psychometrics.

Frequently used data types and machine learning algorithms are introduced. Then, we summarize psych-modeling's application to various scenarios together with its strengths and weaknesses. Compared with traditional self-reporting methods, psych-modeling has some advantages, including better performance in retrospective studies, greater ecological validity, and greater time-efficiency. However, psych-modeling has several limitations. For example, researchers need to spend extra time and effort to learn this new method and bear the inevitable cost of hardware. In future studies, researchers could investigate further how user's behavior on social media relates to psychological indexes. We also expect psych-modeling will be used in future psychological studies. By combining psychometrics and machine learning, we believe psych-modeling could make great contributions to psychology research and practice in the future.

Key words: Psych-Modeling, psychological measurement, social media, machine learning, psychological prediction model